

②

1a. REPORT SECURITY CLASSIFICATION <b>Unclassified</b>			1b. RESTRICTIVE MARKINGS <b>DTIC FILE COPY</b>		
2a. SECURITY CLASSIFICATION <b>AD-A225 998</b>			3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited		
2b. DECLASSIFICATION			4. PERFORMING ORGANIZATION REPORT NUMBER(S)		
5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR- 90 0894</b>			6a. NAME OF PERFORMING ORGANIZATION Howard C. Nusbaum, Dept. of Psychology, University of Chicago		
6b. OFFICE SYMBOL (If applicable) <b>NL</b>			7a. NAME OF MONITORING ORGANIZATION <b>Air Force Office of Scientific Research/NL</b>		
6c. ADDRESS (City, State, and ZIP Code) <b>Chicago, IL 60637</b>			7b. ADDRESS (City, State, and ZIP Code) <b>Building 410 Bolling AFB, Washington, D.C. 20332-6448</b>		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION <b>AFOSR</b>			8b. OFFICE SYMBOL (If applicable) <b>NL</b>		
8c. ADDRESS (City, State, and ZIP Code) <b>Building 410 Bolling AFB, Washington, D.C. 20332-6448</b>			9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <b>AFOSR-89-0389</b>		
10. SOURCE OF FUNDING NUMBERS			11. TITLE (Include Security Classification)		
PROGRAM ELEMENT NO. <b>61102F</b>			PROJECT NO. <b>2313</b>		
TASK NO. <b>A6</b>			WORK UNIT ACCESSION NO. <b>1000 1990</b>		
12. PERSONAL AUTHOR(S) <b>Howard C. Nusbaum Judith C Goodman</b>			13a. TYPE OF REPORT <b>Final Technical Report</b>		
13b. TIME COVERED <b>FROM 15 May 89 to 15 Nov 89</b>			14. DATE OF REPORT (Year, Month, Day) <b>1990, July, 6</b>		
15. PAGE COUNT <b>24</b>			16. SUPPLEMENTARY NOTATION		
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Speech perception, perceptual development, spoken language, auditory perception, word recognition, phonetics.		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Listeners use language-specific knowledge to impose perceptual structure on speech. Since this knowledge must be learned through linguistic experience, data on the development of speech perception could constrain theories of adult speech perception. This workshop focused on the transitions in the perceptual processing of speech from infancy to adulthood. The workshop consisted of nine presentations over a two-day period in June, 1989. Several researchers argued that dynamic mechanisms are critical for relating the acquisition and use of linguistic knowledge to perceptual analysis in a theory of speech perception. Specifically mechanisms of learning and attention were targeted as being critical to the development of speech perception. Further, since listeners appear to use information from several levels of linguistic analysis and to cross-correlate knowledge about perception and production, a complete theory of spoken language understanding must address how different forms of linguistic knowledge are related in perceptual processing.					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION <b>Unclassified</b>		
22a. NAME OF RESPONSIBLE INDIVIDUAL <b>Dr. John F. Tangney</b>			22b. TELEPHONE (Include Area Code) <b>(202) 767-4278</b>		
			22c. OFFICE SYMBOL <b>NL</b>		

24 JUL 1990

For  
X

tion/  
ity Codes  
and/or  
Special

**J. C. Goodman**

Department of Psychology  
University of California, San Diego  
La Jolla, California 92093

Dist		Special	
A-1			



The Workshop on Recognizing Spoken Language was held at the University of Chicago on June 22-23, 1989 specifically to address this issue. The goal of the workshop was to examine transitions in the perceptual processing of speech from infancy to adulthood and to relate perceptual processing to the use of linguistic knowledge. Although there is a substantial body of research concerning the ability of infants to discriminate segmental contrasts in speech, we do not know how to relate these perceptual abilities to the ultimate goal of speech perception, that is the recognition of spoken words in fluent discourse. The decision to focus on this issue in the workshop was based on two assumptions. First, the development of psychological mechanisms is coherent over time: Processes and representations cannot be totally dissociated from their origins. Thus, data concerning developmental changes in the processes involved in spoken language comprehension might constrain theories of speech perception by adults. A complete theory of speech perception must be able to account for data at any point in development. Theories of speech perception should integrate the findings of research concerning language processing across periods of development, rather than account for the perceptual abilities of a single age group. This has not

generally been the case; rather, the development of speech perception has been viewed as a series of discrete stages.

Our second assumption was that the processes and representations of linguistic knowledge are coherent across levels of analysis. Any language is a hierarchical *system* for communication. For example, the representation of spoken words in the lexicon is systematically related to the phonological rules that govern the patterns and transformations of sound sequences in a language. As a result, lexical knowledge may interact with phonological knowledge in speech perception. Data concerning the processes involved in understanding linguistic information at one level of representation may constrain the types of processes that operate at other levels of representation. A complete theory of speech perception should integrate the findings of research concerning language processing across levels of perceptuo-linguistic analysis. However, most current theories treat each level (e.g., phoneme identification and word recognition) as unrelated issues.

Consider some of the current approaches to speech perception. For example, theories of infant speech perception have been very narrow in scope. They have focused on innate perceptual mechanisms, in particular, on the discrimination of phonetic contrasts. Far less work has been directed at characterizing how these perceptual mechanisms (which can be accounted for along general auditory as opposed to language-specific dimensions) come to support recognition of linguistic categories, such as phonemes and words. However, understanding how language-specific knowledge is acquired may well constrain our notions of how listeners impose structure on a stream of speech.

Similarly, there have been very few attempts to relate young children's abilities to perceive segments, as well as to learn and recognize words, to the earlier perceptual abilities of infants or to the process of spoken word recognition in adults. A thorough examination of this transitional period could suggest the nature of mechanisms responsible for spoken language understanding and how these mechanisms change with development. In short, it is necessary to integrate the findings and conclusions of research with infants, young children, and adults in order to construct a complete theory of speech perception.

The primary goal of this workshop was to examine transitions in the perceptual processing of speech from infancy to adulthood. The workshop participants were scientists who have carried out speech research with infants, children, and/or adults. Their task was to consider in detail the theoretical implications of their research for a complete theory of speech perception - in particular, to speculate about how their findings constrain the nature of the mechanisms and representations that mediate speech perception during infancy, childhood, and adulthood. The talks are currently being edited for a book that summarizes the findings of research on speech perception across developmental periods and provides hypotheses of the theoretical principles that govern speech perception and its development.

## **II. Structure and Goals of the Workshop**

We asked workshop participants to apply their empirical findings to basic theoretical questions and to consider the constraints their work places on theories of speech perception. Our broad goal was to provide a discussion and conceptual integration of critical issues in the study of speech perception and the development of this ability. In discussing the goals of the workshop with the invited speakers, we focused on four specific questions: (1) What constraints are placed upon a complete theory of speech perception in order to account for the abilities and knowledge of language users across different stages of development? (2) Is it necessary to propose both a theory of speech perception and a theory of perceptual development or can a single integrated theory account for the development of spoken language comprehension? (3) How are the perceptual mechanisms and linguistic representations proposed in theories of adult speech perception constrained by the infant's sensory abilities? (4) How are segmental perception and word perception related at different stages of development?

Nine research groups presented their empirical work and theoretical perspectives during the two-day workshop. These talks lasted about an hour and were followed by about a half an hour of discussion. In addition, following the last talk, a summary discussion session was held in which the theoretical issues and questions raised during the workshop presentations, as well as suggestions for new directions for perceptual research, were examined in greater detail.

The proceedings of the workshop will be published by MIT Press. Each participant is currently preparing a chapter. Two additional speech scientists are preparing chapters as well; these researchers had been invited to speak at the workshop, but were unable to attend. The discussions that took place during the workshop were recorded and are being used to write a final chapter for the book. We have three goals in writing this chapter. First, we hope to present a conceptual integration of the different perspectives represented by the participants. Second, we wish to present a coherent description of how speech perception changes over time. Third, we plan to present new hypotheses and suggest the types of data that could test these hypotheses.

## **III. Summary of Theoretical Issues**

Abstracts of the talks and the two additional invited chapters are presented in Section V. Each of the presentations raised a number of theoretical issues related to our initial four questions. Over the course of the presentations and discussion, it began to appear as if there is reasonably good agreement on some of the theoretical issues that must be addressed in answering these questions. In particular, six basic theoretical issues were raised across the various discussions that appear to be critical to constructing complete theories of speech perception: (1) the role of learning in speech perception, even beyond its role in development, (2) the role of attention in speech perception, (3) the role of the structure of the lexicon and its relationship to other forms of linguistic knowledge, (4) the role of segmentation of linguistic units in recognition, (5) the role of stress and prosody in speech perception, and (6) the role of relations

between perception and production in the development and use of linguistic knowledge.

Current theories of speech perception cannot account for the development of perceptual abilities because they do not include a mechanism of change as part of the perceptual process. However, a number of workshop participants demonstrated that experience with language modifies perceptual processing both for children and, more significantly, for adults. Thus, dynamic mechanisms are critical for theories concerning the recognition and comprehension of spoken language. Several talks suggested that understanding these dynamic mechanisms may constrain theories of adult speech perception.

Infants are capable of discriminating a universal inventory of phonetic contrasts. These discriminations appear to be based on general auditory, rather than speech-specific, mechanisms. Thus, although the infant is equipped with a set of speech-relevant auditory abilities, these abilities must become attuned to the particular properties of the child's native language. Werker shows that during the first year of life a reorganization begins to occur, such that listeners apparently lose sensitivity to contrasts that are not used phonemically in their native languages. Best shows that this is due to a gradual attentional shift resulting from the initial acquisition of a phonological system. Initial studies suggested that the loss of a phonological contrast during language acquisition was a permanent change due to loss of unused phonetic feature detectors. However, Pisoni, et al. present several studies showing that listeners can be trained to relearn nonnative contrasts. Through experience with a language, listeners learn to focus attention to phonemically-relevant aspects of the signal.

There appear to be three general processes that are important to learning the phonological and phonetic properties of one's native language. First, Best and Studdert-Kennedy argued separately for the importance of experience *producing* spoken language on the organization of perceptual knowledge of phonology. Research reported by Margoliash on perceptual learning in birds supports this general view that experience producing communicative signals develops neurological structures that ultimately may mediate perception of those signals. Further support for this position comes from Mayberry's work on the perceptual processing of American Sign Language. Her research indicates that a critical learning period exists during which specialized neurological structures may be established for perceptual processing of language. The development of these structures is based on early experience producing and perceiving pattern-meaning relations. Even when groups of signers are matched in years of experience, those who learned while young that such relations exist (whether through sign language or spoken English) are able to comprehend language much more effectively than those first exposed to a systematic language later in life.

The second process responsible for phonological organization and structure appears to be the acquisition of a lexicon. As knowledge of the sound patterns of individual spoken words grows, there seems to be a change in the perceptual processing of speech. Jusczyk proposes that infants initially represent words with relatively undifferentiated internal structure, but as vocabulary size grows,

the sound patterns of words are reorganized within the lexicon resulting in more detailed phonological representations. In Jusczyk's theory, listeners weight information according to correlations of distributional properties that signal meaningful distinctions (for example, phonemes distinguish words). This general view of the impact of the growth of the lexicon on phonological knowledge is also supported by Best, Werker, Studdert-Kennedy, Mayberry, Gerken, and Pisoni in their own empirical studies.

Finally, some evidence exists suggesting that experience processing fluent speech—i.e., whole utterances rather than isolated words—is important for the development of phonological knowledge and perceptual processing. For example, Cutler shows that segmentation strategies for isolating linguistic units within the stream of speech depend on linguistic experience. Listeners learn the structural properties of the sound patterns of their language and exploit regularities to find word boundaries. Thus, different languages induce different perceptual segmentation strategies in native listeners based on the rhythmic properties of the speech. Also, Gerken discusses how children and adults differ in their use of the metrical structure of speech. Although there are differences in the production of the stress patterns of speech with age, children do perceive and use the unstressed function words and closed class morphemes of English. She argues that these syntactically important morphemes play a central role in understanding spoken language, even if their representation may be somewhat different in children and adults. These studies suggest the importance of the prosodic properties of whole utterances in learning about and processing the sound patterns of spoken language.

In short, listeners learn about the perceptual properties of linguistic categories through specific experiences using their native language. This knowledge is then used during recognition to impose linguistic structure on the speech signal. Nusbaum and Goodman propose that the imposition of structure through linguistic knowledge is carried out by shifting and organizing perceptual attention to the acoustic properties of the speech signal. They argue that the development of speech perception in children and recognition of spoken language in adults are mediated by the same general cognitive mechanisms. Adults learn to focus attention on specific properties of the speech signal to make linguistic judgments using the same kinds of mechanisms and information that young children use. Furthermore, these mechanisms can operate across different levels of processing allowing listeners to learn the constraints of a particular talker and context that facilitate recognition of phonemes or words or prosodic structure. This introduces the theoretical construct of constraint satisfaction as a means of identifying linguistically relevant units from an utterance — that is, information from a variety of sources (acoustic-phonetic, semantic, syntactic) is used together to converge on the identity of a segment or a word given that no single source is sufficient on its own. The nature of constraining information may change with experience with one's language, but the mechanisms that learn and apply this information do not need to change.

This view of speech perception incorporates dynamic mechanisms that account for the way speech perception develops throughout childhood and, at the same time account for the way speech is perceived by adults whose behavior was

once thought to be static. In other words, theories of speech perception need not be distinct from theories that explain perceptuo-linguistic development. If adult speech perception is mediated by shifting attention to perceptual constraints based on learning constraints of particular context and talker characteristics, then there should be similarities in the ability of adults and children to learn new properties of speech.

Indeed, several workshop participants noted parallels between perceptual abilities of infants and adults, suggesting that a unified account of spoken language comprehension should be constructed. However, there is little agreement about the way to explain these similarities in perceptual performance. For example, Miller argues that mechanisms responsible for similarities in infant and adult speech perception are the result of innate speech-specific processors. However, she does not propose any developmental mechanisms that can account for the differences in performance as a function of age. Thus, Miller's view is that similarities in infant and adult perception are explained by common mechanisms, but these mechanisms are not the same as those responsible for development.

Other participants propose more general mechanisms that allow shifts in the nature of information that is most constraining at different points in development. Pisoni, et al. note that during development listeners shift their attention to different aspects of the sound patterns of speech, resulting in changes to their memory representations for distinctive phonemes; these representations are established by listeners' experiences. Similarly, Nusbaum and Goodman cite evidence of learning by adults based on experience as well as by children. They propose that listeners learn which linguistic categories are relevant to signalling meaning distinctions. Further, listeners are able to shift their attention between levels of linguistic function and use information from several linguistic levels simultaneously to comprehend speech. Along these lines, Gerken presents data that young children are learning word- and sentence-level constraints simultaneously, while others have argued that children learn phonological and word level constraints simultaneously. With experience, knowledge of the sorts of information that occur together to constrain the potential identity of various linguistic categories may change, but the process of analyzing the signal does not. Further, constraints are drawn from more than one level of linguistic category. In sum, processes per se do not change, but the knowledge with which a listener interprets a speech signal does change with development and experience.

If this is true, then it is important to understand how spoken language is represented mentally and how representations at different levels of linguistic abstraction interact with each other and with the analysis of the signal during speech perception for both adults and children. Traditionally, most theories of speech perception have focused on explaining a specific perceptual phenomenon at a single level of representation. For example, theories of phoneme perception are directed at explaining how listeners recognize phonetic information given the variability in the acoustic properties of phonemes produced in different phonetic contexts and produced by different talkers. These theories focus primarily on the bottom-up processing of the acoustic properties of speech ending at the level of the phoneme.

In contrast with this traditional approach, two presentations use developmental data to suggest that listeners distribute attention across several levels of linguistic representation at one time and that information across representations is used constrain recognition at any particular level. Gerken shows that very young children attend simultaneously to word- and to sentence-level information, and knowledge of word-level information constrains production as well. Nusbaum and Goodman show that very young children simultaneously attend to sentence-level semantic context and to acoustic-phonetic information when perceiving words. Thus, very young children are able to use multiple sources of information to constrain recognition of linguistic categories in much the same way as adults. This suggests that the perceptual mechanisms that integrate information for recognizing speech are very similar in adults and children. The changes that result during the development of speech perception are not due to changes in the underlying processing mechanisms; instead these changes appear to result from increased knowledge about the structural properties of spoken language (Gerken; Mayberry; Nusbaum and Goodman; Pisoni). Further, as a result of linguistic experience, listeners assign relative weights to various sources of information; during speech perception, the focus and distribution of attention at any one point in time depends on this knowledge (Jusczyk; Nusbaum and Goodman; Pisoni, Logan, & Lively). The effect of experience with one's native language is to learn how to direct attention to the acoustic, segmental, lexical, and sentential properties of an utterance that work together to specify the linguistic interpretation of the utterance. Listeners learn how various levels of linguistic representation are coordinated within their linguistic system and as a consequence they direct attention away from those properties that are not linguistically relevant or informative. The outcome of this learning is that adults lose the ability to make phonetic contrasts that do not signal meaning differences in their native language even though they could discriminate those perceptual distinctions as infants (Best; Werker). However, appropriate training can redirect attention back to those properties and recover the ability to make those contrasts (Pisoni, et al.). Thus, theories of speech perception need to incorporate mechanisms of learning and attention, although none currently do.

#### **IV. Conclusions**

Speech research has all too often separated the experiments on infant speech perception from experiments on adult speech perception. Data from infants generally have been compared with data from adults but treated as if they are reflective of precisely the same perceptual system (e.g., Miller) or very different perceptual systems (e.g., Mayberry). Instead, we propose that the data from infants may constrain theories of adult speech perception. Similarly research on perception of phonemes has generally been treated as disjoint from the research on perception of spoken words or sentences. Some researchers have focused on the former level of linguistic representation (e.g., Best), others on the latter (e.g., Mayberry). We suggest interactive processing of different perceptual-linguistic levels.



One important aspect of the majority of presentations and discussions at the workshop was the acknowledgement that it is important to begin to consider how the adult perceptual system emerges from the infant's initial abilities and how different levels of linguistic knowledge interact. Perhaps the strongest organizing theme for integrating these areas of research was the view that learning and attention—general cognitive mechanisms—may be critical to developing theories of speech perception. A reasonable working assumption is that the perceptual mechanisms of the infant are retained intact in the adult. What changes is knowledge about the linguistic structures of one's native language. As the child gains more experience perceiving and producing spoken language, more is learned about how language is a coherent system whose parts and levels of representation are mutually constraining. In this regard then, it becomes important to investigate speech perception in children throughout a range of ages beyond the first year. In the past, there has been a tacit assumption that it was possible to study infants and adults and extrapolate development as a direct line between them. If the interesting questions in the development of speech perception concern the acquisition of linguistic knowledge, how the process of acquisition affects the ongoing use of that knowledge, and how different types of knowledge interact, it becomes difficult to interpolate directly between an infant with minimal linguistic knowledge to the adult with relatively complete knowledge.

Children learn how talkers differ in producing speech, how phonemes interact in their acoustic realizations and how stress affects the phonological representations of words. For any particular utterance, adults must learn about the characteristics of the specific talker that produced it and the specific context in which it is produced in order to constrain the recognition and comprehension of that utterance. Children learn to direct attention to sets of linguistically relevant properties of utterances and for any particular utterance, adults must learn which of the known constraints will be effective. In order to understand how listeners at any age process and recognize spoken language, it will be important to investigate how general cognitive mechanisms such as learning and attention are used in speech perception.

The presentations and discussions at the workshop have made it clear that in order to understand the transitions that occur throughout the development of speech perception, we need to examine how children learn the structural properties of their native language and how they exploit this developing knowledge. Speech perception cannot be understood by restricting investigations to phoneme perception or word perception or to infants or to adults. Moreover it appears as though it will no longer be sufficient to explain speech perception by alluding to specialized perceptual mechanisms and ignoring the role of more general cognitive and perceptual processes. By examining and systematically relating the perceptual abilities of adults and children, as they operate within the context of a rich system of linguistic knowledge using general cognitive mechanisms, we may be able to make substantially more progress in understanding speech perception than has been made in the last 40 years. In this way, the papers that will be published from this workshop may foreshadow or perhaps even set the research agenda for speech science for the coming years.

**V. Abstracts**

*Sentential Processes in Early Language:  
Evidence from the Perception and Production of Function Morphemes*

LouAnn Gerken  
State University of New York at Buffalo

This chapter questions the view that referential content words are the building blocks for children's early sentences by demonstrating their sensitivity to function morphemes, such as articles and verb inflections. Because function morphemes serve as markers to the syntactic form of sentences, the degree to which children demonstrate tacit knowledge of these elements is one measure of their knowledge of sentence-level processes. In particular, data will be presented indicating that children attend to function morphemes when listening to sentences and that they use function morphemes in sentence comprehension. In addition, it appears that children's representation of function morphemes in sentences is very similar to their representation of content words, suggesting that word level and sentence level processes overlap greatly in early language.

*Context-Sensitivity in Phoneme Perception*

Joanne Miller  
Northeastern University

A complete theory of speech perception must account not only for the perceptual abilities of adults, but also for the origins of such abilities in young infants and the transition between infant and adult perception over the course of language acquisition. Over the past two decades, it has become increasingly clear that infants come to the task of language acquisition with highly sophisticated processing systems for speech. The implications of this finding for theories of adult speech perception, and the nature of the interplay between infant and adult data vis a vis theories of speech perception, are discussed in terms of one well-studied phenomenon, the context-sensitivity of phonetic perception.

*Cross-language speech perception:  
Developmental change does not involve loss.*

Janet Werker  
University of British Columbia

There are profound changes across age in cross-language speech perception performance, but an adequate understanding of how and why these developmental changes take place is still lacking. In this chapter, our research in cross language speech perception will be reviewed, with special reference to how it fits into the short "history" of research in cross-language speech perception. In this endeavor, three overlapping "periods" of research in cross-language speech perception will be differentiated. The first period includes the early seminal work indicating that adults have difficulty both perceiving and producing some non-native phonetic contrasts, but that infants can apparently discriminate non-native contrasts with ease. The research findings from this period culminated in the hypothesis that age-related changes in cross-language speech perception resulted from an a sensory "loss". The second period, lasting from is characterized by a rising skepticism with respect to the adequacy of the "loss" explanation. Research conducted during this period indicated that adults can be trained to discriminate non-native contrasts, that some non-native contrasts can be discriminated without any training, and that there are significant differences in results depending upon the testing procedure employed. The third period is characterized by a firm understanding that "loss" is not an adequate explanation for age- and experiential influences on cross-language speech perception, and that more complex explanations need to be articulated. This search for more adequate explanations has led to a rich proliferation of theoretically motivated research.

*Developments in infants' perception of nonnative speech contrasts: Transitions from phonetic universals to language-specific phonology*

Catherine Best  
Wesleyan University and Haskins Laboratories

Both general observation and empirical studies indicate that linguistic experience affects the perception of nonnative phonetic contrasts by adults. That is, mature listeners often have some difficulty discriminating speech contrasts that are not employed in their familiar native language(s). However, research with young infants has shown more universal speech perception abilities, such that most native and nonnative contrasts alike are discriminated and there appears to be no clear influence of the language environment on this ability until sometime during the second half-year of life. Thus, an effect of linguistic experience is that discrimination of at least some phonetic contrasts declines when these are not utilized in the native language. This effect is apparently attentional, rather than a permanent sensory-neural change, and some findings suggest that it is related to the phonemic rather than the phonetic or auditory levels of perception. Recent studies with infants indicate that this perceptual reorganization has already begun to take place by about 8-10 months of age, suggesting that the infant may begin shifting to a phonological, or phonemic, level of perceiving speech at that time.

The experiments reported at the conference were designed to further explore the possible phonological basis of the emergence of language-specific influences on infant speech perception. We began with the general hypothesis that for listeners operating at a phonemic level, nonnative phones become perceptually assimilated to the most similar phoneme category in the native language. In this model, similarity is defined in terms of articulatory-gestural properties rather than purely acoustic ones. The hypothesis leads to four predictions about discriminability of various types of nonnative contrasts: 1) a contrast will be difficult to discriminate if both nonnative phones assimilate equally well to a single native category (SC); 2) it will be significantly easier to discriminate if one nonnative phone assimilates better than the other to the same native category, thus showing a category goodness difference (CG); 3) discrimination will be very good if the nonnative phones assimilate to two different native categories (TC); 4) discrimination will be good if the nonnative phones are not assimilable (NA) to native categories and hence are treated as nonspeech. We tested these predictions with adult English speakers, using contrasts of each assimilation types from the Zulu language; the pattern of predictions was upheld by their performance (TC > CG = NA > SC). We then also tested infants at 6-8 months and at 10-12 months, since previous research had shown language-specific influences by the older age but not the younger. As predicted, the younger infants discriminated across all types of nonnative contrasts. As expected, the older infants did differ from the younger ones in discrimination performance. However, they did not show the same pattern of discrimination as the adults had, suggesting that they were not yet perceiving at a phonological level. Although they discriminated the NA contrast well and the SC

contrast poorly, like the adults, they failed to discriminate the TC and CG contrasts on which the adults had performed quite well.

On the basis of these results, I suggest that development in speech perception proceeds from a universal sensitivity to phonetic-articulatory properties in early infancy, through a period of dawning recognition of some aspects of gestural combinations found in native phones beginning in the last quarter of the first year, to mature perception of the gestural details of phonemic categories and contrasts by adulthood. This approach may also help to explain important aspects in the early development of productive phonology, as illustrated in additional data from my laboratory on a child's acquisition of one complex phonological rule in spoken English.

*Perceptual learning of nonnative speech contrasts:  
Implications for theories of speech perception*

David Pisoni, John Logan, and Scott Lively  
Indiana University

The goal of this chapter is to examine the effect of early linguistic experience on the speech perception abilities of adults. It has long been believed one's native linguistic environment produces permanent changes in identification and discrimination of speech sounds. Thus, infants appear to perceive a wide variety of nonnative speech contrasts, while adults only seem to perceive those in their native languages. In fact, efforts to selectively modify speech perception abilities using laboratory training techniques have generally met with little success. In this chapter, we present two experiments in which listeners were successfully trained to perceive a nonnative contrast. In the first experiment, native speakers of English were trained to perceive prevoiced stimuli. In the second experiment, native speakers of Japanese were trained to perceive the /r/-/l/ contrast. Contrary to prior findings, both experiments demonstrated that the underlying sensory-perceptual mechanisms have not been permanently modified or lost by prior linguistic experience. Methodological differences account for the discrepancy between earlier findings and the present results.

These findings suggest that current theories of speech perception must be modified to allow for perceptual change. We propose that linguistic experience affects perception by modifying attentional processes. One consequence of this view of selective attention for speech perception is that the relative salience of different phonetically relevant dimensions depends on the language-learning environment. Thus, similarity relations for different perceptual dimensions are tied to linguistic experience. A second consequence is that there are changes in the memory representations for psychologically more salient dimensions. This view of the role of selective attention in speech perception can accommodate a wide variety of developmental and cross-language findings in the literature and provide a psychological basis for the mechanisms underlying perceptual change.

*Infant Speech Perception and the Development of the Mental Lexicon*

Peter W. Jusczyk  
University of Oregon

How does the ability to locate and identify words in a stream of speech emerge? The development of a number of functions that are prerequisites for word recognition are discussed. First, listeners must be capable of discriminating utterances of one word type from those of another (for example, 'bat' vs. 'pat'). Second, they must be capable of categorizing correctly all the different tokens of the same utterance type. For example, they must ignore the acoustic differences that arise in uttering the same word from one occasion to another, or by one speaker or another. Third, listeners must be able to locate words from a stream of speech. That is, they must segment the signal into word-sized units. Fourth, listeners must represent the sound pattern of a word along with its meaning in memory. The representation must be distinctive enough to differentiate it from similar-sounding words during speech recognition. Consequently, an important part of the recognition process depends on how the sound properties of the word are encoded in memory, since incoming speech signals must be matched against stored perceptual representations to be recognized and for their meanings to be recovered. Finally, the role of attentional and memory processes in speech perception must be considered. Specifically, to what aspects of the speech signal do infants attend, and what information about the speech signal is retained in memory?

A model of the development of word recognition is proposed which is based on the assumption that a listener extracts an array of properties which are then weighted according to their importance in signaling meaningful distinctions in the language. The mastery of the sound structure of a language entails acquiring the appropriate weighting scheme, which is established by attending to correlations of distributional properties in a particular language. The resulting weighted representation is matched against lexical representations stored in memory. As a listener gains experience with a language, his weighted analysis of the speech signal will constrain the nature of information about words that is retained in memory.



*Contributions of the Study of Birdsong to Speech Science.*

Dan Margoliash  
University of Chicago

Neuroethology grew in response to the inability of behavioral techniques to address mechanistic questions that could resolve problems of behavioral process. Since language has uniquely human qualities, speech science cannot directly embrace reductionistic approaches such as single cell physiology. The study of birdsong learning has generated much interest among speech scientists because it exhibits behavioral constraints during development similar to speech yet is amenable to the neuroethological approach. This chapter reviews the current theoretical status of aspects of birdsong learning that have also captured attention in the speech world. These include an assessment of similarities and differences in the two developmental phenomena, the feature detector hypothesis, relationships between production and perception in birdsong, a more general discussion of 'special' mechanisms, and recent work on adult vocal plasticity in birds. Although song learning can only serve as a limited model of speech acquisition, nevertheless there have been useful exchanges of ideas and concepts, and there appears to be considerable room for continued interactions in the future.

*Marked and Unmarked Segmentation Strategies?*

Anne Cutler  
MRC Applied Psychology Unit

The goal of this chapter is to examine how listeners deal with the segmentation problem in continuous speech. The segmentation problem refers to the processes by which a listener divides the speech signal into discrete linguistic units such as words. Since the acoustic waveform rarely contains reliable cues to the presence of a word boundary, listeners must apply strategies to the utterance to achieve maximum segmentation efficiency. We report several experiments that address two specific questions. First, are the processes listeners use to segment the speech signal universal or language-specific? Second, if strategies are language-specific, are they properties of listeners or of languages? That is, can any listener employ the strategy appropriate to the language with which he or she is presented, or are large amounts of experience with the language necessary. We hypothesized that the segmentation strategy for English is based on stress rhythm, while the segmentation strategy for French is syllable-based. These hypotheses were based on specific structural properties of each language, so the strategies are unlikely to be universal. Through a series of syllable-monitoring tasks with bilingual English/French speakers, we were able to determine that some segmentation strategies are universal (or unmarked) while others require early experience with one's native language.

*Learning to Hear Speech as Spoken Language*

Howard C. Nusbaum  
The University of Chicago

and Judith C. Goodman  
The University of California, San Diego

Theories of speech perception have traditionally sought to explain the abilities of adult listeners without considering the genesis and development of those abilities. As a result, each of these theories really engenders two separate theories: an explicit perceptual theory and some unspecified theory of perceptual development. Research on the development of speech perception has not altered this separation in any substantive way. Indeed it has been shaped by our understanding of adult speech perception, without regard for the process of development. As with adults, research on infant speech perception has been concerned primarily with the development of segmental contrasts, although only the simplest of such questions have been asked. These studies have been interpreted as suggesting that perceptual development is relatively complete at the end of the first year of life. Unfortunately this has led to the simplest possible theories of perceptual development; these data have not suggested any constraints on theories of adult speech perception.

Speech perception involves more than segmental discrimination based on the most rudimentary acoustic properties. Comprehension of spoken language depends on much more elaborate segmental mechanisms, along with a broad range of abilities to recognize, understand, and integrate other acoustic and linguistic patterns in speech. The time course of development for these more complex abilities is not complete within a single year of exposure to language. As research on the development of speech perception moves to consider these more sophisticated perceptuo-linguistic abilities, these data will have stronger implications for theories of adult speech perception.

In contrast to the currently accepted view that separates theories of adult speech perception from theories of the development of that ability, we propose that a single theory should account both for speech perception by adults and for development. According to this perspective, theories of adult speech perception should have a very different form from current theories, and developmental data should be very relevant to testing those theories. In order for a theory of speech perception to account for its own development, it must include a mechanism of change as part of the perceptual process. Since current theories of speech perception do not include dynamic mechanisms, there is no straightforward way to modify them to accomplish this goal. A theory of speech perception that can account for its own development may better explain adult perception as well: There are data that suggest that even for adults, dynamic mechanisms are a critical part of the processes that mediate recognition and comprehension of spoken language.

Our view is that speech perception and its development depend on three theoretical constructs: learning, attention, and constraint satisfaction. From

this perspective, speech perception is not a simple recognition process that maps acoustic properties onto linguistic categories and development is not a simple process of determining which mappings are appropriate for a particular language. The problem of lack of invariance between pattern structure and linguistic interpretations in speech argues against that kind of approach. Instead, speech perception may be a process of constraint satisfaction in which linguistic categories are inferred according to criteria that are based on mental theories of the linguistic function of those categories. The specific constraints that specify category membership may shift with those factors that are responsible for the lack of invariance in speech and so listeners must dynamically shift attention from one set of constraints to another, using the same mechanisms by which these constraints are learned during childhood.

*The Importance of Childhood to Language Acquisition:  
Insights from American Sign Language*

Rachel Mayberry  
McGill University

The circumstances under which deaf people acquire sign language are remarkably heterogeneous. Some deaf signers first acquire sign language during early childhood in situations analogous to native language acquisition. Other deaf signers first acquire sign language after childhood in circumstances that are analogous to second-language acquisition. However, some deaf signers first acquire sign language after childhood in circumstances that are distinct from those of either native or second-language acquisition -- a situation best characterized as late first-language acquisition. What are the consequences of these unique and diverse circumstances of sign language acquisition on its long-range outcome? In this paper we summarize five studies designed to answer the question. The results show that age of acquisition exerts robust effects on deaf signers' ability to process and comprehend sign language, even after 42 years or more of continuous practice. The effects of acquiring sign language after childhood are greater for first-language acquisition than second-language acquisition. Lexical processing and narrative comprehension are especially affected, but effects are present at nearly every level of psycholinguistic analysis. The overall pattern of results suggests that the young child is better able to induce and master phonological systems than the older child, adolescent, or adult.

*Understanding adults' speech structure through research on child speech*

Michael Studdert-Kennedy  
Haskins Laboratories

[NOTE: Studdert-Kennedy has not complied with repeated requests for either an abstract of his presentation or for his chapter.]

**VI. Participants in the Workshop on Recognizing Spoken Language**

**Speakers:**

Howard Nusbaum and Judith Goodman  
The University of Chicago  
*Learning to hear speech as spoken language*

Michael Studdert-Kennedy  
Haskins Laboratories  
*Understanding adults' speech structure through research on child speech*

David Pisoni, John Logan, and Scott Lively  
Indiana University  
*Perceptual learning of non-native phonetic contrasts:  
Some new data on /r/ and /l/*

Catherine Best  
Haskins Laboratories  
*Developments in infants' perception of nonnative speech contrasts: Transitions  
from phonetic universals to language-specific phonology*

Anne Cutler  
*Marked and unmarked segmentation strategies?*

Peter Jusczyk  
University of Oregon  
*Infant speech perception and the development of a mental lexicon*

LouAnn Gerken  
SUNY/Buffalo  
*Perceiving the closed class*

Rachel Mayberry  
McGill University  
*The importance of childhood to language acquisition:  
Insights from American Sign Language*

Dan Margoliash  
The University of Chicago  
*How theories of birdsong can inform speech scientists and other fantasies*

**Attendees:**

Kevin Broihier, University of Chicago  
Starkey Duncan, University of Chicago  
Susan Duncan, University of Chicago  
Anne Farley, University of Chicago  
Steve Goldinger, Indiana University  
John Goldsmith, University of Chicago  
Jenny de Groot, University of Chicago  
Beth Greene, Indiana University  
Janellen Huttenlocher, University of Chicago  
Karen Landahl, University of Chicago  
Lisa Lee, University of Chicago  
Susan Levine, University of Chicago  
Jerre Levy, University of Chicago  
David McNeill, University of Chicago  
Todd Morin, University of Chicago  
Nancy Stein, University of Chicago